Appendix 1: Baseflow and Baseflow Metrics

Contents

A.1 Baseflow Separation Techniques

A.1.1 HYSEP

The following description of HYSEP is taken from (Schwartz 2007):

(Sloto, Crouse and Geological Survey (U.S.) 1996) describe three algorithms to automate heuristic baseflow separation from streamflow records. The USGS HYSEP program automates the 1) fixed block; 2) sliding block; and 3) local minimum methods. These algorithms use the empirical relationship between basin drainage area and the characteristic response time $N = A^{0.2}$. N is the estimated duration of surface runoff after the hydrograph peak in days, for basin drainage area A in square miles (Chow et al. 1988, Linsley 1958 , Viessman et al. 1977). Baseflow is estimated using data windows of length τ, and the half window of length $t' = (\tau-1)/2$, where $\tau = 2N^*$ days and $2N^*$ is defined as the odd integer closest to 2N, s.t. $3 \leq 2N^* \leq 11$.

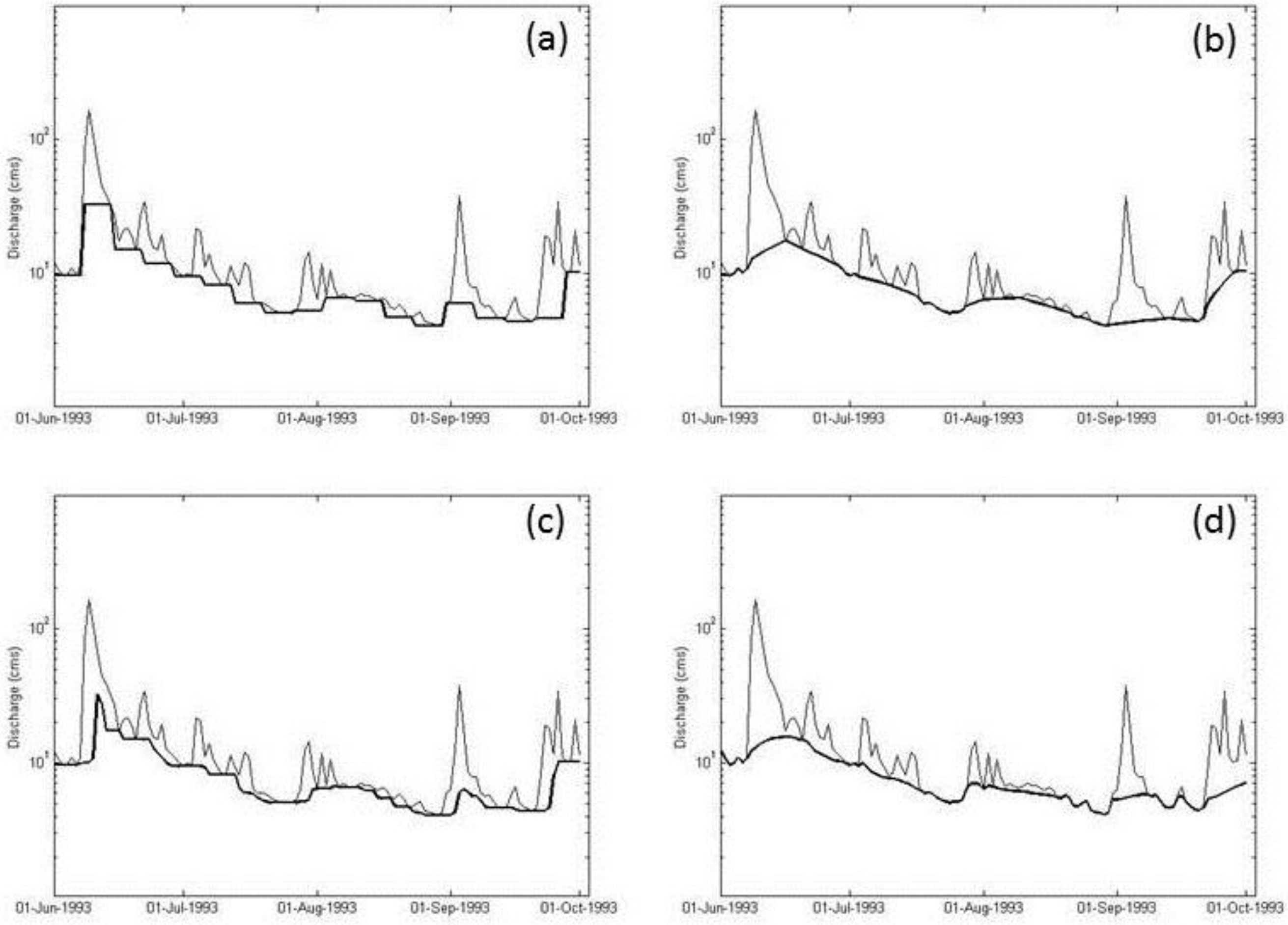


Figure 1 - Baseflow Separation for USGS gauge 04208000 Using (a) Fixed Block; (b) Local Minimum; (c) Sliding Block; (d) Digital Filter: Third Pass

Fixed Block Method The fixed block method starts on day t_0, and defines the baseflow for τ consecutive days (t_0, t_0+1, ... $t_0+\tau-1$) as $q_b(t_0) = \min\{Q_t,\ t \in [\ t_0,\ t_0+\tau-1]\ \}$. Here t_0 is the first day of the n-th block of τ days, for $t_0 = n\tau+1$ where $n = \mathrm{int}[t/\tau]$, and $\mathrm{int}[t/\tau]$ denotes the integer part of t/τ , $t = 1,2, ..., T$.

<u>Sliding block Method</u> The sliding block method defines baseflow on each day t as the minimum flow in the τ day window centered on day t; $q_b(t) = \min\{Q_t,\ t \in [\,t-t', t+t'\,]\}$ $t = t'+1, t'+2, \ldots, T-t'$.

<u>Local Minimum Method</u> The local minimum method begins by constructing a sequence of "local minima". The discharge on day t, Q_t, is defined as a local minimum if it is the lowest observed flow in the interval [t-t', t+t']. Each local minimum is considered the baseflow value on that day. Baseflow on all other days is computed through linear interpolation between successive minima.

A.1.2 PART

PART is a streamflow partitioning program developed by the USGS to estimate a daily record of baseflow (Rutledge 1998). Like the HYSEP algorithm, PART implements a characteristic response time in days that is a function of the drainage area of the watershed in square miles, $N = A^{0.2}$ where N is rounded down to the nearest integer. N is the estimated time until the end of surface runoff from a storm event, so if daily streamflow values decrease for N or greater days, then the antecedent recession requirement is met and it is assumed that baseflow is equal to streamflow for these days. PART performs a linear interpolation between every day in the record that meets this requirement, with an error check to ensure that baseflow never exceeds streamflow and that streamflow never decreases by more than 0.1 log cycle on days where flow is assumed to come entirely from groundwater. Groundwater discharge is summarized for the entire period of record using an antecedent recession requirement of N, N+1, and N+2 days, and the program makes a second order polynomial for groundwater discharge as a function of this antecedent recession requirement. Baseflow is evaluated for the exact value of N from $N=A^{0.2}$ using this polynomial.

MatLab code was written to perform this method of heuristic baseflow separation on vectors of daily discharge with date numbers, and results from the USGS program matched results from the MatLab program.

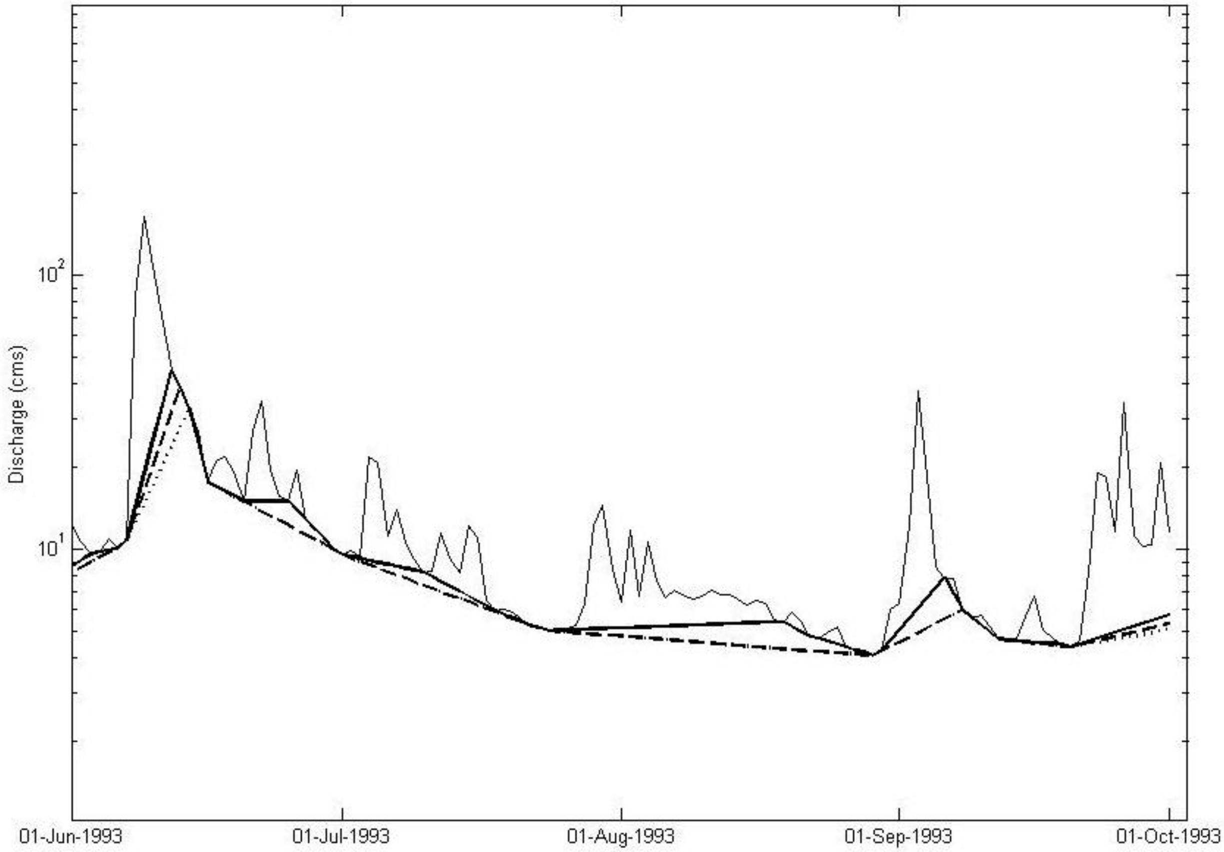


Figure 2 - PART baseflow separation on USGS gauge 04208000. The solid, dashed and dotted lines represent baseflow separation using antecedent recession requirements of N, N+1, and N+2 days respectively.

A.1.3 Digital Filter

The digital filter described in (Nathan and McMahon 1990) was implemented in MatLab. Vectors of time and daily discharge are imported into MatLab, which runs multiple passes of the digital filter. The algorithm for the filter is:

$$f_k = \alpha f_{k-1} + \frac{(1+\alpha)}{2}(y_k - y_{k-1})$$

$$b_k = y_k - f_k \qquad \text{provided that} \qquad 0 \le b_k \le y_k$$

Where y_k is the discharge, f_k is the filtered quickflow, and b_k is the baseflow. This filter is run three times (the second pass runs backward from the final observation to the first). Here and throughout we will use the third pass of the digital filter as a consistent estimator of daily baseflow as suggested by Nathan and McMahon.

The five different methods of baseflow separation yield five different estimates of baseflow index, summarized in the table below for the HCDN gauges used in this study

Table 1 - Comparing Baseflow Index using different baseflow separation algorithms

Gauge Name	USGS Gauge Numer	Water Year POR	DIGITAL FILTER BFI	HYSEP FIX INT BFI	HYSEP SLIDING INT BFI	HYSEP LOCAL MIN BFI	PART BFI
DeepCreekMannboro	02041000	1947-1988	0.494	0.544	0.547	0.513	0.543
AppomattoxFarmville	02039500	1927-1988	0.550	0.551	0.549	0.535	0.581
HolidayCreekAnderson	02038850	1967-1988	0.570	0.664	0.658	0.610	0.643
JohnsCreekNewCastle	02017500	1927-1988	0.488	0.603	0.603	0.544	0.653
BuffaloRiverTye	02027800	1961-1988	0.615	0.695	0.696	0.659	0.729
CraigCreekParr	02018000	1926-1988	0.498	0.544	0.547	0.514	0.623
JamesRiverCartersville	02035000	1925-1979	0.586	0.566	0.568	0.554	0.596
DragonSwamp	01669500	1944-1981	0.596	0.733	0.731	0.644	0.747
SlateRiverArvonia	02030500	1937-1988	0.529	0.560	0.560	0.536	0.561
PottsCreekCovington	02014000	1929-1956	0.508	0.617	0.612	0.564	0.636
PottsCreekCovington	02014000	1966-1988	0.507	0.615	0.613	0.546	0.643
CowPasture	02016000	1926-1988	0.498	0.525	0.528	0.503	0.576
DunlapCovington	02013000	1929-1988	0.454	0.538	0.537	0.494	0.553
HardwareRiver	02030000	1939-1988	0.595	0.652	0.650	0.621	0.679
CalfPasture	02020500	1939-1988	0.407	0.514	0.511	0.452	0.522
MattaponiBowling	01674000	1943-1988	0.504	0.544	0.545	0.524	0.570
BullPasture	02015700	1961-1988	0.546	0.630	0.630	0.588	0.665
StMarys	01661500	1947-1988	0.474	0.608	0.608	0.533	0.547
Rappahannock Fredericksburg	01668000	1911-1988	0.535	0.529	0.525	0.511	0.565
RobinsonLocust	01666500	1944-1988	0.591	0.664	0.663	0.620	0.695
StClements	01661050	1969-1988	0.509	0.604	0.605	0.553	0.567
RapidanCulpeper	01667500	1931-1988	0.574	0.606	0.605	0.575	0.661
Rappahannock Remington	01664000	1943-1988	0.554	0.584	0.588	0.563	0.635
HazelRixeyville	01663500	1943-1988	0.564	0.601	0.603	0.571	0.677
ShenandoahCootes	01632000	1926-1988	0.362	0.484	0.480	0.410	0.474
Piscataway	01653600	1966-1988	0.479	0.520	0.521	0.502	0.555
Nanticoke	01487000	1944-1969	0.689	0.787	0.788	0.748	0.823
Beaverdam	01492000	1951-1981	0.404	0.514	0.510	0.451	0.476
MarshyHope	01488500	1944-1968	0.503	0.594	0.600	0.552	0.619
ShenandoahFrontRoyal	01631000	1961-1988	0.599	0.627	0.627	0.601	0.693
PotomacAdjustedDC	01646502	1931-1988	0.582	0.584	0.584	0.561	0.588
ShenandoahStrasburg	01634000	1926-1988	0.554	0.608	0.609	0.572	0.655
NorthRiver	01590000	1933-1974	0.699	0.775	0.775	0.738	0.771
Choptank	01491000	1949-1983	0.546	0.624	0.627	0.591	0.636
GooseCreekLeesburg	01644000	1931-1988	0.499	0.547	0.544	0.508	0.609

CedarCreek	01634500	1938-1988	0.475	0.580	0.578	0.527	0.611
SenecaDawsonville	01645000	1931-1975	0.621	0.659	0.659	0.641	0.687
PatuxentUnity	01591000	1945-1988	0.638	0.682	0.683	0.663	0.731
bennett	01643500	1967-1988	0.588	0.641	0.643	0.613	0.682
NorthPotomacSteyer	01595000	1957-1988	0.500	0.600	0.600	0.566	0.629
CatoctinMiddletown	01637500	1948-1988	0.523	0.642	0.640	0.586	0.703
PatapscoNBranch	01586500	1928-1953	0.677	0.712	0.711	0.693	0.739
PattersonHeadsville	01604500	1939-1988	0.423	0.556	0.555	0.491	0.567
SouthPotomac Springfield	01608500	1929-1988	0.502	0.520	0.520	0.504	0.577
GeorgesCreekFranklin	01599000	1930-1977	0.496	0.642	0.643	0.585	0.689
SladeGlyndon	01583000	1948-1981	0.711	0.785	0.786	0.768	0.806
Crabtree	01597000	1949-1981	0.478	0.743	0.741	0.612	0.675
WesternRun	01583500	1945-1988	0.720	0.753	0.754	0.740	0.786
BackCreekJones	1614000	1939-1975	0.416	0.532	0.529	0.479	0.528
MarshRunGrimes	01617800	1964-1988	0.758	0.891	0.891	0.865	0.898
FishingCreekLewistown	01641500	1948-1984	0.648	0.874	0.872	0.818	0.892
PotomacPawPaw	01610000	1939-1980	0.523	0.507	0.509	0.496	0.557
CacaponRiverWV	01611500	1924-1988	0.470	0.529	0.528	0.497	0.569
LittleFallsBlueMount	01582000	1945-1988	0.731	0.764	0.764	0.750	0.799
BigPipeBruceville	01639500	1948-1988	0.569	0.616	0.617	0.594	0.644
Principio	01496200	1968-1988	0.548	0.603	0.601	0.572	0.596
DeerCreekRocks	01580000	1927-1988	0.709	0.739	0.737	0.723	0.763
BigElkCreek	01495000	1933-1988	0.626	0.635	0.637	0.627	0.666
WillsCreekCumberland	01601500	1930-1988	0.486	0.552	0.550	0.511	0.624
OwensCreek	01640500	1932-1984	0.553	0.752	0.750	0.688	0.758
MonocacyBridgeport	01639000	1943-1988	0.336	0.386	0.390	0.369	0.380
PotomacHancock	01613000	1933-1988	0.518	0.506	0.506	0.486	0.554
ConococheagueFairview	01614500	1929-1988	0.563	0.623	0.624	0.588	0.681
EvittsCreek	01603500	1933-1982	0.514	0.729	0.729	0.643	0.694
DunningCreekBelden	01560000	1940-1988	0.465	0.591	0.589	0.536	0.581
WestConewago	01574000	1929-1959	0.427	0.443	0.440	0.419	0.472
AughwickCreek	01564500	1939-1988	0.438	0.557	0.557	0.502	0.561
RaystownBranchJuniata	01562000	1912-1988	0.503	0.562	0.565	0.527	0.594
ShermanCreek	01568000	1930-1988	0.503	0.616	0.615	0.573	0.636
FrankstownBranch Juniata	01556000	1917-1988	0.548	0.599	0.599	0.571	0.655
JuniataNewport	01567000	1901-1972	0.555	0.552	0.552	0.533	0.604
MahantangoCreek	01555500	1930-1988	0.501	0.627	0.625	0.566	0.659
LittleJuniataSpruce	01558000	1939-1988	0.592	0.720	0.719	0.669	0.746
PennsCreek	01555000	1930-1988	0.581	0.677	0.677	0.640	0.758
WestBranch Susquehanna	01541000	1914-1988	0.481	0.523	0.524	0.493	0.575
ClearfieldCreekDimeling	01541500	1914-1960	0.486	0.541	0.542	0.510	0.586
Wapwallopen	01538000	1920-1988	0.553	0.672	0.671	0.628	0.735

MarshCreekBlanchard	01547700	1956-1988	0.462	0.622	0.623	0.553	0.673
FishingCreekBloomsburg	01539000	1939-1988	0.511	0.581	0.583	0.554	0.664
Sinnemahoning	01543500	1939-1988	0.474	0.547	0.549	0.518	0.585
YoungWomansCreek	01545600	1966-1988	0.511	0.696	0.695	0.602	0.763
BlockhouseCreek	01549500	1941-1988	0.458	0.597	0.597	0.533	0.655
PineCreekCedar	01548500	1919-1988	0.488	0.587	0.585	0.547	0.635
TunkhannockCreek	01534000	1915-1988	0.471	0.529	0.529	0.507	0.579
TowandaCreek	01532000	1915-1988	0.420	0.539	0.540	0.502	0.537
TiogaRiver	01518000	1939-1977	0.446	0.508	0.510	0.483	0.565
CowanesqueRiver	01520000	1952-1979	0.401	0.460	0.458	0.422	0.506
ChemungRiver	01531000	1916-1948	0.439	0.458	0.459	0.453	0.446
TiogaLindley	01520500	1931-1978	0.426	0.491	0.489	0.462	0.502
SusquehannaConklin	01503000	1914-1988	0.542	0.602	0.603	0.577	0.621
OwegoCreek	01514000	1931-1977	0.451	0.594	0.592	0.556	0.589
ChenangoRiver	01512500	1914-1941	0.517	0.561	0.567	0.556	0.591
FiveMileCreek	01528000	1938-1988	0.411	0.543	0.546	0.510	0.554
CharlotteCreek	01498500	1939-1975	0.516	0.669	0.670	0.644	0.687
ButternutCreek	01502000	1939-1988	0.517	0.661	0.663	0.630	0.707

A.2 Estimating Recharge

A.2.1 Baseflow and the Water Balance

The simplest estimate of recharge is simply baseflow. Under a stationary water balance, it is assumed there is no change in groundwater storage and that groundwater inflow and outflow are equal. Neglecting evapotranspiration, lateral flow and interbasin transfers, long term recharge is approximately equal to long term baseflow. Using long records of gauged streamflow data and heuristic baseflow separation techniques yields estimates of these components in the water balance.

A.2.2 RECESS

The original RECESS program was an interactive process to select periods of recession that are linear or nearly linear in log space (Rutledge 1998). The program scanned a record of daily discharge values to determine if each day was a peak [Q(i) > Q(i-1) and Q(i) > Q(i+1)] and define the length of recession after each peak. If the recession meets the user specified minimum length (Rutledge suggests using 10 days), the user has the option to retain the recession after visually assessing its linearity in a plot of time from peak vs. the log of discharge. For each retained recession, linear regression is performed for log of flow as a function of time. The negative of the slope of this regression line is defined as recession index K, which is the number of days for flow to decrease by one log cycle. Averaging these K values from each recession yields the mean value of the recess index for the gauge.

Instead of the time consuming interactive process the USGS RECESS program, a MatLab implementation was developed to automate the selection of linear recessions using user defined parameters. This MatLab program selects recessions like the USGS program, by identifying each peak and the subsequent recession and omitting N days after each peak. Instead of manually choosing which of these recessions to retain, a linear regression is performed on all recessions that meet the user defined minimum length requirement. The R^2 statistic of each regression is determined, and only recessions with an R^2 above the user defined threshold are retained for use in determining the recession index K. The tables below compare K values from the MatLab program's to results from the USGS Regional Aquifer Systems Analysis on the Appalachian Valley and Ridge (Rutledge, Mesko and Geological 1996). It is difficult to reproduce exactly the K values from the user implementation of the USGS RECESS program; even repeated analyses of the same gauge by the same user can return different estimates of K. Rutledge acknowledges this and suggests that a single user use the program when comparing K values across different gauges. To avoid variation from user selection of recessions, the MatLab implementation uses a consistent set of user defined parameters instead of interactive selection to ensure that the recession index is determined in the same manner across all gauges. Suggested parameters for the MatLab program are highlighted in the tables below, as determined by an analysis of recession indices across all gauges in the RASA study.

USGS Gauge 02041000 for calendar year 1947-1991			
User Defined Min. Recession	User Defined R^2 Theshold	# Recessions	Mean K
8	0.8	151	52.5
8	0.85	143	51.9
8	0.9	134	51.5
8	0.95	90	48.8
8	0.99	14	43.6
10	0.8	79	55.9
10	0.85	73	55
10	0.9	68	53.8
10	0.95	48	49.4
10	0.99	5	56.2
12	0.8	44	54.7
12	**0.85**	**41**	**54.2**
12	0.9	37	51.4
12	0.95	27	47.5
12	0.99	2	50.8
USGS RASA Results			
-	-	23	58.1

Table 2 - Comparison of MatLab program and USGS RASA reported Recession index K

Values of recession index K are used in the determining groundwater recharge with the RORA program and a change in the K value will lead to a change in estimated recharge. To evaluate the automated RECESS algorithm, the sensitivity of RORA's recharge estimates to recession index K was investigated.

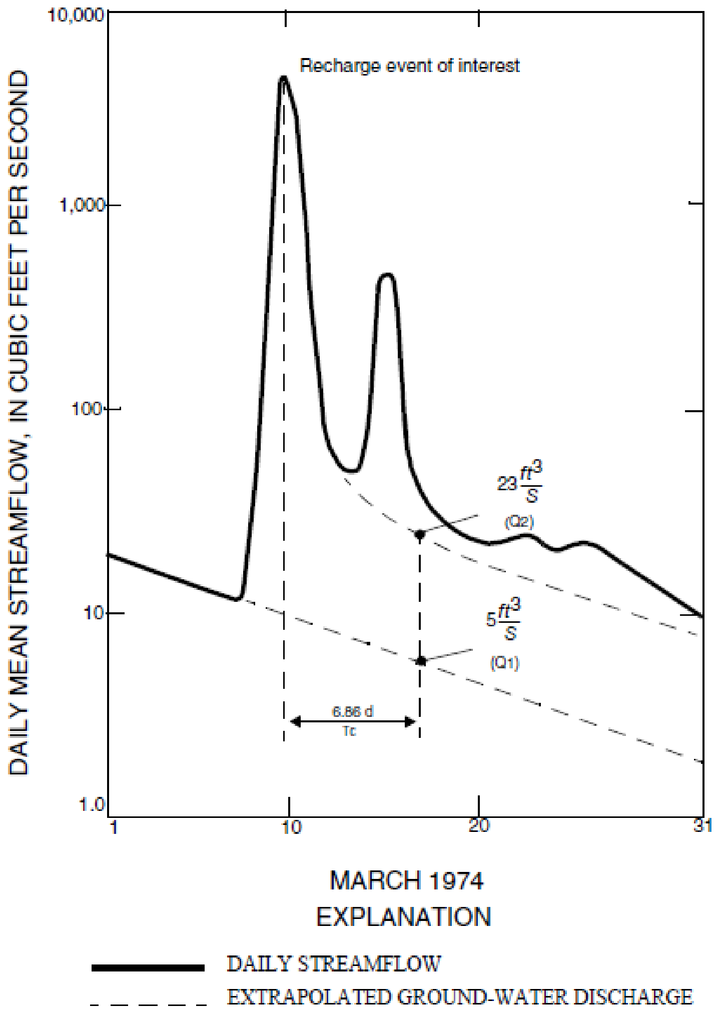
Table 3 below shows that recharge estimates are insensitive to variations in K. K can vary from 20 to 60 (a 300% change) while the estimate of mean annual recharge only varies from 10 to 10.3 (3% change). The estimates of recession index K from the MatLab implementation will be acceptable for use in RORA for determining groundwater recharge.

Little Patuxent at Guilford **1/1/1933 to 12/31/2009**		
K value	Total Recharge (in)	Mean Annual Recharge (in)
20	794.3	10.3
25	798.3	10.4
30	796.7	10.3
35	795.2	10.3
40	789.0	10.2
45	782.9	10.2
50	777.2	10.1
55	772.0	10.0
60	767.2	10.0

Table 3 - Sensitivity of RORA Recharge estimates to Recession Index K

A.2.3 RORA

The USGS original program RORA implements recession curve displacement to provide an estimate of groundwater recharge using a time series of daily stream discharge data (Rutledge 1998). The program identifies peaks in streamflow and the recession period in between each peak and compares extrapolated baseflow estimates from before and after each peak to provide an estimate of recharge from that storm event. RORA uses a characteristic response time of N days, a function of the drainage area in square miles, where $N = A^{0.2}$ and is rounded up to the nearest integer. This is the number of days that streamflow must be decreasing to be considered a recession. The critical time T_C (the number of days after a peak to which baseflow is extrapolated) is a function of recession index K taken from RECESS. A peak is defined as the largest daily value of streamflow between two recession periods. RORA extrapolates a hypothetical baseflow recession from before the storm event using the baseflow recession index K to a point T_C days after the peak in streamflow. Baseflow is similarly extrapolated from the recession period following the storm event. These two extrapolated baseflow recessions are compared at T_C days after peak, and this difference is directly related to the change in groundwater storage and provides a consistent estimate of net recharge from a storm event. Recharge estimates are calculated over long periods of record and can easily be summarized as average annual recharge in inches. The algorithms from the USGS program were coded into MatLab, and the recharge estimates from the original program were identical to the results from the MatLab code.



PROCEDURE

1. Compute recession index, K (32d/ log cycle)
2. Compute critical time, Tc (0.2144xK, or 6.86 d)
3. Locate time that is 6.86 days after peak
4. Extrapolate pre-event recession to Q1 (5 ft^3/s)
5. Extrapolate post-event recession to Q2 (23 ft^3/s)
6. Compute total recharge,

$$\frac{2x\left(18\frac{ft^3}{s}\right)x32d}{2.3026}x\frac{86400s}{1d} = 4.32x10^7 ft^3$$

Figure 3 - Figure taken from (Rutledge 1998)

A.3 Hydraulic Baseflow Indices

A.3.1 Baseflow Recession Constant K_b

While baseflow separation techniques are a useful and simple way to estimate the recharge component of the water balance that can account for hydrologic variation, they do not account for hydraulic variation within the watershed. The water balance is modulated by both changes in hydrologic fluxes (e.g. precipitation, ET) and soil and aquifer characteristics such as hydraulic conductivity and the volume and release rate of stored groundwater – the dimensionless storatvity. Using a conceptual model of recharge and discharge from an idealized aquifer, the "effective" characteristics of aquifer systems may be inferred and estimated from the characteristics of observed baseflow through recession analysis of gauged streamflow (Vogel and Kroll 1996). Using the Dupuit- Bousinessq representation of a saturated aquifer draining to a fully penetrating stream, the recession constant K_b can be related to basin-scale effective aquifer characteristics(Vogel and Kroll 1992). For unimpaired recharge, we expect K_b to provide a regional signature of effective basin characteristics such as soil porosity and hydraulic conductivity. Physical properties of regional flow systems such as hydraulic conductivity and thickness and extent of unconfined aquifers are not expected to change with development, thus significant changes in streamflow response are inferred to result from changes in forcings (e.g. infiltration, runoff, ET) representing the baseflow signature of development.

Kroll's algorithms were implemented in an original MatLab program that accepts a timeseries of daily streamflow measurements and determines the baseflow recession constant. Recession segments are selected from the hydrograph using criteria: "A recession begins when a 3 day moving average begins to decrease and ends when the average begins to increase." Any recessions shorter than the minimum length of ten days are discarded. For the remaining recession, the first portion of 0.3λ days from a λ day recession is also discarded. Then, for each pair of discharge data within these selected recessions (pair meaning discharge on day *t* and day *t+1*), the following equation is applied to return a Kb_5 value for the gauge.

$$exp\left\{-\exp\left[\frac{1}{m}\sum_{t=1}^{m}\left\{\ln[Q_{t-1} - Q_t] - \ln[\frac{1}{2}(Q_t + Q_{t+1})]\right\}\right]\right\}$$

When the baseflow component of streamflow accounts for all the discharge in a stream, consider this portion of the hydrograph to be a "baseflow recession." The percentage of discharge that carries over to the next day is an approximation of the baseflow recession constant Kb. A lower value of Kb means that baseflow recedes at a faster rate, or that less baseflow persists each day.

Nathan, R. J. & T. A. McMahon (1990) Evaluation of automated techniques for base flow and recession analyses. *Water Resour. Res.,* 26**,** 1465-1473.

Rutledge, A. T. 1998. Computer Programs for Describing the Recession of Ground-Water Discharge and for Estimating Mean Ground-Water Recharge and Discharge from Streamflow Records - Update. In *Water-Resources Investigations Report*. Reston, Virginia: U.S. Geological Survey.

Rutledge, A. T., T. O. Mesko & S. Geological. 1996. *Estimated hydrologic characteristics of shallow aquifer systems in the Valley and Ridge, the Blue Ridge, and the Piedmont physiographic provinces based on analysis of streamflow recession and base flow*. Washington; Denver, CO: U.S. G.P.O. ; For sale by U.S. Geological Survey, Information Services.

Schwartz, S. (2007) Automated algorithms for heuristic base-flow separation. *Journal of the American Water Resources Association***,** 1583-1594.

Sloto, R. A., M. Y. Crouse & Geological Survey (U.S.). 1996. *HYSEP, a computer program for streamflow hydrograph separation and analysis*. U.S. Dept. of the Interior

Branch of Information Services [distributor]. http://purl.access.gpo.gov/GPO/LPS96165 (last accessed.

Vogel, R. M. & C. N. Kroll (1992) Regional geohydrologic-geomorphic relationships for the estimation of low-flow statistics. *Water Resour. Res.,* 28**,** 2451-2458.

--- (1996) Estimation of baseflow recession constants. *Water Resources Management,* 10**,** 303-320.